Web Design & Programming

Hosting

Xavier Belanger

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

http://creativecommons.org/licenses/by-sa/4.0/

You are free to:

- **Share** — copy and redistribute the material in any medium or format
- **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

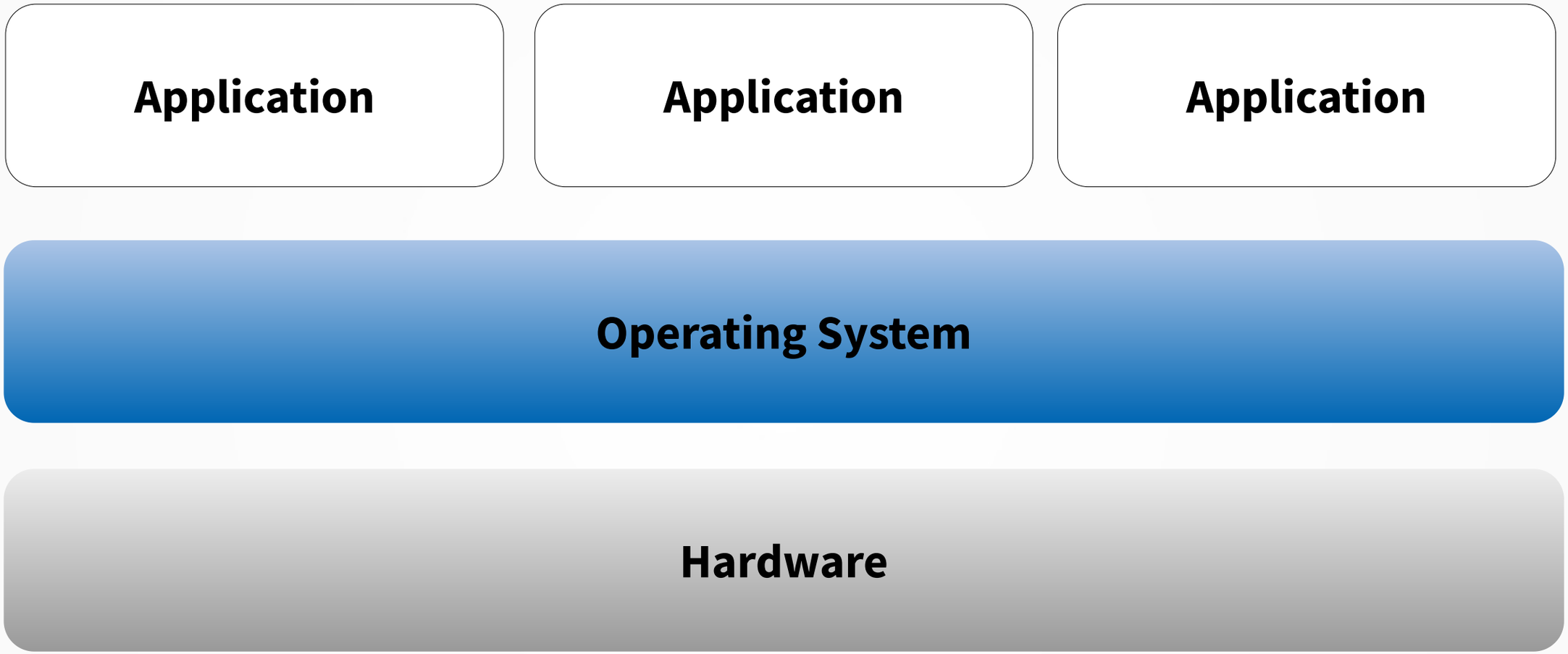
- **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.
- **No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Where to host a server?

- Physical Machine (bare metal)
- Virtual Machine
- Containers
- Cloud

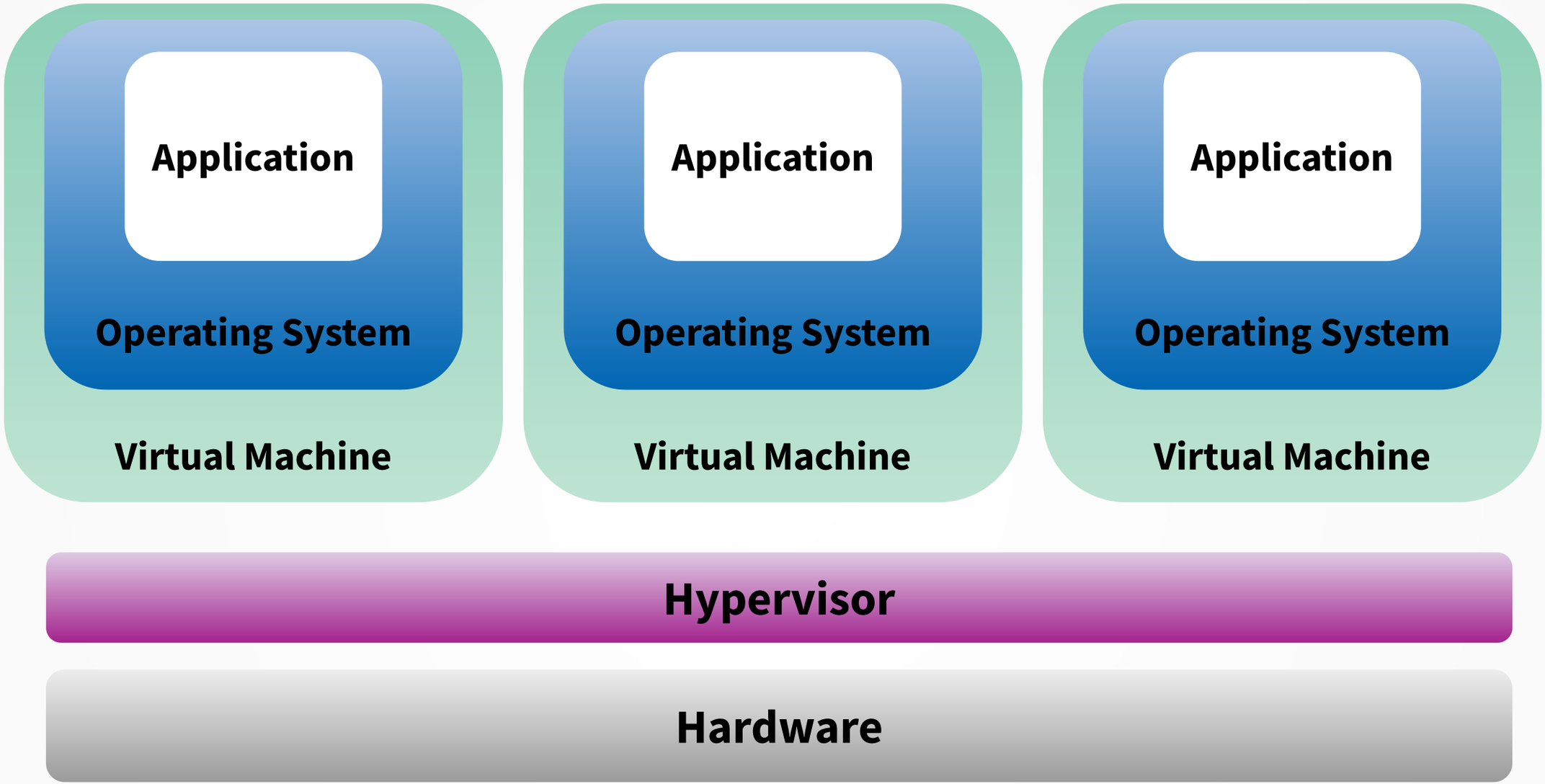
Physical Machine

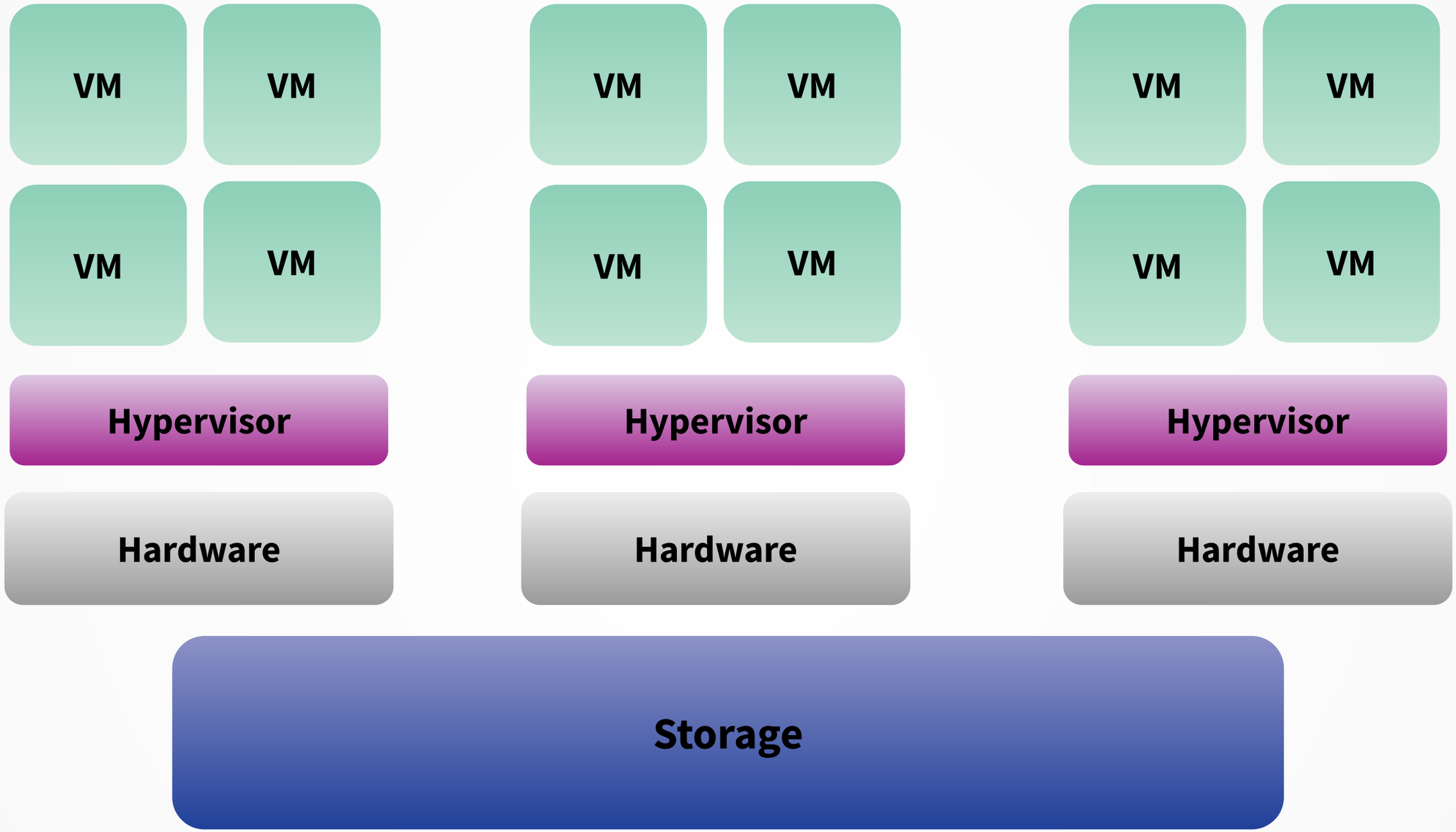
- Expensive.
- Require hardware maintenance and support contract.
- Can be hosted on premises, or in a data center (colocation).
- Less and less common.


Application
Application
Application
Operating System
Hardware

Virtual Machine

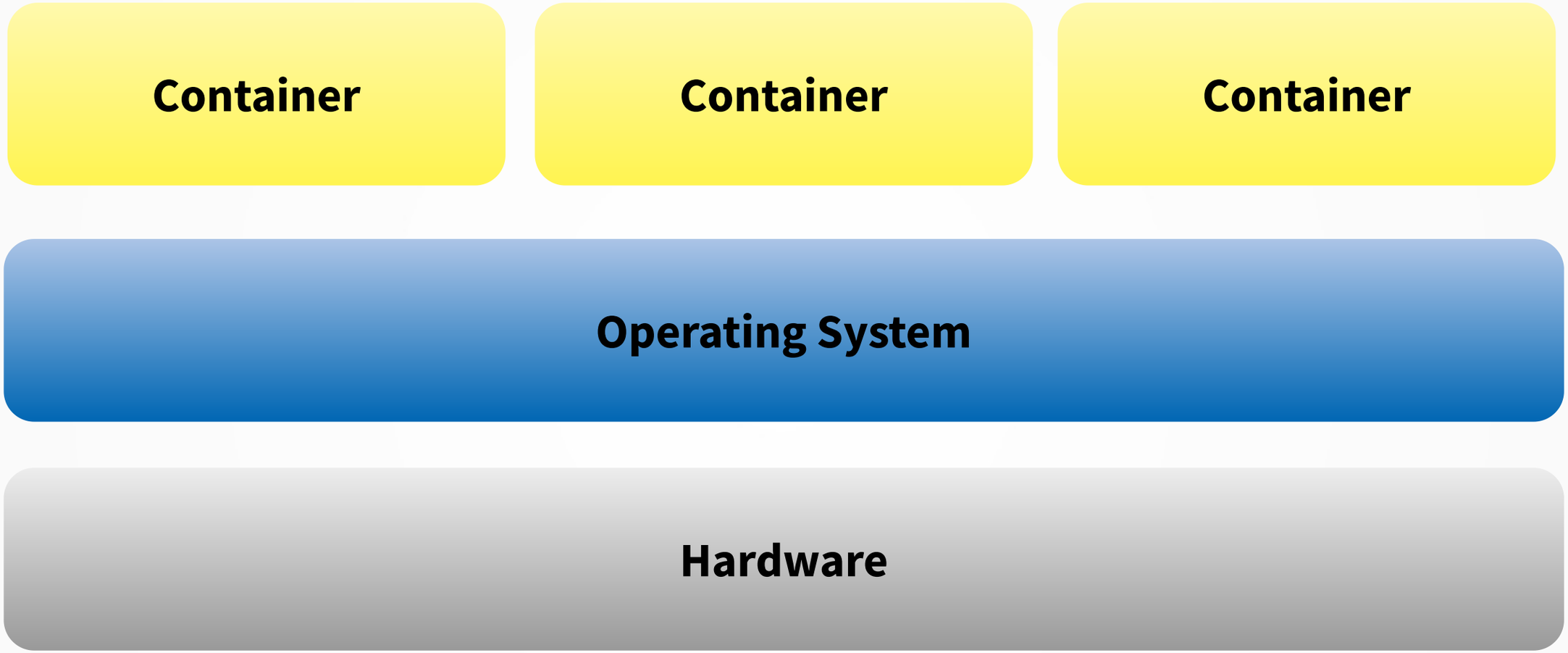
- Probably the most common solution as of today.
- Can be used as a simple “standalone” system or as a complex, redundant hosting architecture.
- Special features: snapshots, migration, failover.
- A physical server can be converted to a virtual machine.
  A virtual machine can be cloned to create similar systems.
- Major vendors: VMware vCenter/ESXi, Microsoft Hyper-V, Xen Project / Citrix Hypervisor.


Application
Operating System
Virtual Machine
Application
Operating System
Virtual Machine
Application
Operating System
Virtual Machine
Hypervisor
Hardware


VM
VM
VM
VM
VM
VM
VM
VM
VM
VM
VM
VM
Hypervisor
Hypervisor
Hypervisor
Hardware
Hardware
Hardware
Storage

Containers

- Used mostly for a specific, limited tasks (microservices).
- Containers are often used in numbers.
- Disposable.
- Most common vendor: Docker/Kubernetes.


Container
Container
Container
Operating System
Hardware

Cloud

- Multiple solutions are available, and can be combined.
- Hourly, monthly or yearly billing cycle.
- Major vendors:
  - Amazon Web Services (AWS)
  - Google Cloud Platform (GCP)
  - Microsoft Azure

*aaS

- “Something as a Service”
  - Infrastructure (IaaS)
  - Platform (PaaS)
  - Software (SaaS)
  - and others …

Responsibilities

- Depending on the hosting solution and the vendors, you must pay attention on who is responsible for:
  - System upgrades and patching
  - Network access and security
  - Application upgrades (including languages)
  - Data security and backups (files and databases)
  - User access and permissions (identity management)
  - Monitoring and logging

Web Hosting Providers

- Multiple companies offer specialized web hosting services.
- Other services may be included (DNS, email, monitoring, backups, …).
- Configuration and options may be limited.

Availability and High Availability

- In order to provide access to a web server all components must be up and running: web server, database server, network access, …
- Each component should be monitored and trigger an alert if the service is degraded.
- High Availability (HA) is an architecture using a combination of solutions to limit the impact of one element (or more) being unavailable.

“Nines”

Availability is sometimes expressed as a percentage, giving the value (or setting an objective) of availability during a specified period of time.

Percentage	Maximum downtime in a year
99%	3 days, 15 hours, 36 minutes
99.9%	8 hours, 45 minutes
99.99%	52 minutes, 34 seconds
99.999%	5 minutes, 15 seconds

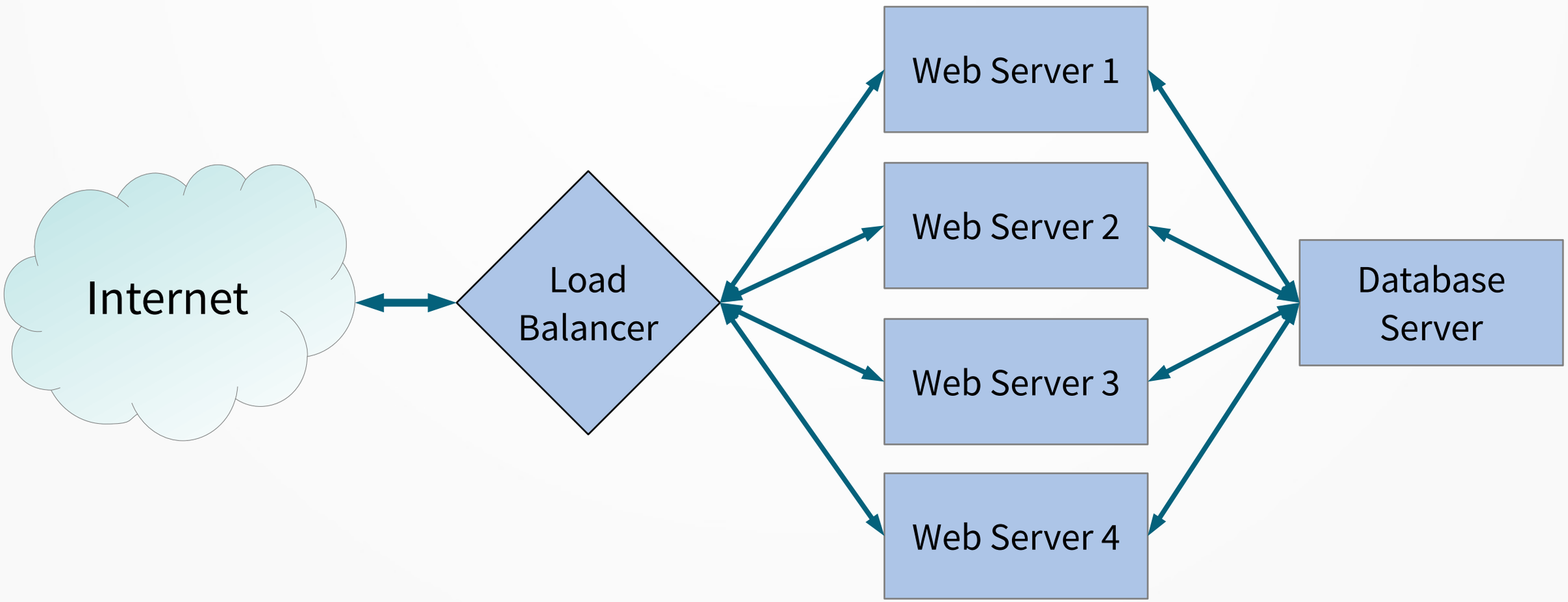
Availability Issues

- Hardware failure
- Service failure
- System failure
- Maintenance (scheduled or unscheduled)
- Traffic increase
- Malicious activity
- Human error
- Natural disaster
- …

Load Balancers

- A load balancer is a device receiving all incoming traffic for a specific service, then the workload is divided between multiple backend nodes.
- Nodes should be identical in configuration (not necessarily in resources).
- Some applications may require some fine-tuned configuration to work with a load balancer (user session tracking for instance).
- Load balancers can also provide additional functions (encryption endpoint / offloading , port redirection, …).

Basic Load Balancer Architecture



Elasticity

- Using cloud provider services it is possible to use load balancers, to create and delete virtual machines or containers, to assign more resources and as a result to scale a service up or down, as needed.
  This is known as “elastic computing”.
- The process must be automatic.
- This can also add complexity with monitoring tools; time constraints should also be taken into consideration.

Content Delivery Network

- A Content Delivery Network (CDN) is a geographically distributed network, responding to requests from memory when possible to improve performance.
- CDN can be specialized for specific type of assets (images, video, …).
- CDN can also provide additional features, like DDoS (Distributed Denial of Service) protection.